



**PCT**  
WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<b>(51) Internationale Patentklassifikation <sup>6</sup> :</b>  <b>C12Q 1/68</b>	<b>A1</b>	<b>(11) Internationale Veröffentlichungsnummer:</b> <b>WO 99/42610</b>  <b>(43) Internationales Veröffentlichungsdatum:</b> 26. August 1999 (26.08.99)		
<table style="width: 100%; border: none;"><tr><td style="width: 50%; vertical-align: top; padding: 5px;"><b>(21) Internationales Aktenzeichen:</b> PCT/EP99/00992  <b>(22) Internationales Anmeldedatum:</b> 16. Februar 1999 (16.02.99)  <b>(30) Prioritätsdaten:</b> 198 06 431.4      17. Februar 1998 (17.02.98)      DE  <b>(71) Anmelder (für alle Bestimmungsstaaten ausser AT US):</b> NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).  <b>(71) Anmelder (nur für AT):</b> NOVARTIS-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT MBH [AT/AT]; Brunner Strasse 59, A-1235 Vienna (AT).  <b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US):</b> HOFFMANN, Ralf [DE/DE]; Am Mittelberg 5, D-51375 Schlebusch (DE). LÜBBERT, Hermann [DE/DE]; Höhenstrasse 59, D-51381 Lützenkirchen (DE). ZWILLING, Stefan [DE/DE]; Münsters Gässchen 20, D-51375 Schlebusch (DE).  <b>(74) Anwalt:</b> BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent &amp; Trademark Dept., CH-4002 Basel (CH).</td><td style="width: 50%; vertical-align: top; padding: 5px;"><b>(81) Bestimmungsstaaten:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Veröffentlicht</b> <i>Mit internationalem Recherchenbericht.</i> <i>Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i></td></tr></table>			<b>(21) Internationales Aktenzeichen:</b> PCT/EP99/00992  <b>(22) Internationales Anmeldedatum:</b> 16. Februar 1999 (16.02.99)  <b>(30) Prioritätsdaten:</b> 198 06 431.4      17. Februar 1998 (17.02.98)      DE  <b>(71) Anmelder (für alle Bestimmungsstaaten ausser AT US):</b> NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).  <b>(71) Anmelder (nur für AT):</b> NOVARTIS-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT MBH [AT/AT]; Brunner Strasse 59, A-1235 Vienna (AT).  <b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US):</b> HOFFMANN, Ralf [DE/DE]; Am Mittelberg 5, D-51375 Schlebusch (DE). LÜBBERT, Hermann [DE/DE]; Höhenstrasse 59, D-51381 Lützenkirchen (DE). ZWILLING, Stefan [DE/DE]; Münsters Gässchen 20, D-51375 Schlebusch (DE).  <b>(74) Anwalt:</b> BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH-4002 Basel (CH).	<b>(81) Bestimmungsstaaten:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Veröffentlicht</b> <i>Mit internationalem Recherchenbericht.</i> <i>Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>
<b>(21) Internationales Aktenzeichen:</b> PCT/EP99/00992  <b>(22) Internationales Anmeldedatum:</b> 16. Februar 1999 (16.02.99)  <b>(30) Prioritätsdaten:</b> 198 06 431.4      17. Februar 1998 (17.02.98)      DE  <b>(71) Anmelder (für alle Bestimmungsstaaten ausser AT US):</b> NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).  <b>(71) Anmelder (nur für AT):</b> NOVARTIS-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT MBH [AT/AT]; Brunner Strasse 59, A-1235 Vienna (AT).  <b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US):</b> HOFFMANN, Ralf [DE/DE]; Am Mittelberg 5, D-51375 Schlebusch (DE). LÜBBERT, Hermann [DE/DE]; Höhenstrasse 59, D-51381 Lützenkirchen (DE). ZWILLING, Stefan [DE/DE]; Münsters Gässchen 20, D-51375 Schlebusch (DE).  <b>(74) Anwalt:</b> BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH-4002 Basel (CH).	<b>(81) Bestimmungsstaaten:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Veröffentlicht</b> <i>Mit internationalem Recherchenbericht.</i> <i>Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>			
<b>(54) Title:</b> METHODS FOR CHARACTERISING mRNA MOLECULES  <b>(54) Bezeichnung:</b> VERFAHREN ZUR CHARAKTERISIERUNG VON mRNA-MOLEKÜLEN  <b>(57) Abstract</b>  The invention relates to methods for the qualitative and quantitative determination of differentially expressed mRNA molecules. Said methods are used especially to determine if possible all mRNA molecules present in a cell or a tissue, and to compare them with other cells or tissues or with other conditions (stages of disease or development) or stages of treatment for same. The method provided for in the invention therefore makes it possible, for example, to establish a comprehensive map of the different mRNA molecules present in a defined mRNA population and subsequently to use the preferably digital information obtained in this way in data base analyses.  <b>(57) Zusammenfassung</b>  Die vorliegende Erfindung betrifft Verfahren zur qualitativen und quantitativen Erfassung differentiell exprimierter mRNA-Moleküle. Diese Verfahren dienen insbesondere der Erfassung möglichst aller in einer Zelle oder in einem Gewebe vorliegenden mRNA-Moleküle sowie deren Vergleich mit anderen Zellen oder Geweben oder mit bestimmten Zuständen (Krankheits- oder Entwicklungsstadien) oder Behandlungsphasen derselben. Das erfindungsgemässe Verfahren erlaubt somit z.B. die Erstellung eines umfassenden Abbildes der in einer definierten mRNA-Population vorliegenden verschiedenen mRNA-Moleküle sowie die anschliessende Verwendung der hierbei erhaltenen, vorzugsweise digitalen Information im Rahmen von Datenbankanalysen.				

### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

### Verfahren zur Charakterisierung von mRNA-Molekülen

Die vorliegende Erfindung betrifft Verfahren zur qualitativen und quantitativen Erfassung differentiell exprimierter mRNA-Moleküle. Die der Erfindung zugrundeliegende Technik wird nachfolgend kurz mit DEPD (Digital Expression Pattern Display) bezeichnet.

Das Genom höherer Organismen umfasst nach gegenwärtigen Schätzungen etwa 100.000 verschiedene Gene, von denen jedoch nur eine vergleichsweise kleine Anzahl in jeder Zelle eines Organismus exprimiert und damit in Polypeptide und Proteine umgesetzt wird. Es wird davon ausgegangen, dass weitgehend alle Prozesse und Stoffwechselleistungen im Bereich der lebenden Materie davon abhängen, welche Gene zu welchem Zeitpunkt in welchen Geweben an- bzw. abgeschaltet werden. So weisen zahlreiche Befunde darauf hin, dass zelluläre Prozesse wie beispielsweise Homöostase, Reaktionen auf Allergien, Regulation des Zellzyklus, Altern und das Eintreten von Zellen in den programmierten Zelltod (Apoptose) auf der differentiellen Expression bestimmter Gene basieren oder hiermit im Zusammenhang stehen. Sowohl der Verlauf einer normalen Entwicklung als auch die zu Erkrankungen wie z.B. Krebs führenden pathologischen Erscheinungen beruhen im wesentlichen auf Veränderungen in der Genexpression.

Dementsprechend besteht Bedarf an spezifischen Verfahren zur Erfassung unterschiedlich bzw. differentiell exprimierter mRNA-Moleküle, um Unterschiede in der Genexpression im Vergleich zu geeigneten Kontrollen zu identifizieren. Derartige Verfahren wären sowohl diagnostisch als auch im Wege der Evaluierung therapeutischer Targets von grosser Bedeutung.

Das erfindungsgemässe Verfahren dient insbesondere der Erfassung möglichst aller in einer Zelle oder in einem Gewebe vorliegenden mRNA-Moleküle sowie deren Vergleich mit anderen Zellen oder Geweben oder mit bestimmten Zuständen (Krankheits- oder Entwicklungsstadien) oder Behandlungsphasen derselben, und zwar vorzugsweise sowohl in qualitativer als auch in quantitativer Hinsicht. Das erfindungsgemässe Verfahren erlaubt somit z.B. die Erstellung eines umfassenden Abbildes der in einer definierten mRNA-Population vorliegenden verschiedenen mRNA-Moleküle sowie die anschliessende Verwendung der hierbei erhaltenen, vorzugsweise digitalen Information im Rahmen von

Datenbankanalysen. Es ist davon auszugehen, dass in naher Zukunft alle humanen Gensequenzen in entsprechenden Datenbanken zur Verfügung stehen werden. Das vorliegende Verfahren erlaubt deshalb die gesamtheitliche Erfassung und Charakterisierung von zellulären Vorgängen, die sich in spezifischen Expressionsmustern auf der Ebene der mRNA-Populationen widerspiegeln. Hierdurch können beispielsweise Veränderungen im Expressionsmuster einzelner, an einem spezifischen Prozess beteiligter Gene schnell und verlässlich identifiziert werden. Dadurch lassen sich auch neue Substanzwirkziele für pharmazeutische Wirkstoffe definieren. Die erhaltenen Informationen können auch eingesetzt werden, um über nachvollziehbare biochemische Signal- bzw. Synthesewege den kausalen Zusammenhang zwischen bekannten Zielgenen und Zielproteinen herzustellen.

Dem Fachmann auf dem betreffenden Gebiet ist die oben dargelegte Aufgabenstellung bekannt. Im Stand der Technik sind zur Lösung der gestellten Aufgabe verschiedene Wege beschritten worden.

Beispielsweise beschreiben P. Liang und A.B. Pardee ein Verfahren zur Trennung individueller mRNAs mittels Polymerase-Kettenreaktion (PCR) (P. Liang & A.B. Pardee 1992 Science 257, 967-971). Dieses Verfahren wurde angewendet, um die von zwei verwandten Zelltypen exprimierten mRNA-Populationen zu vergleichen. Eine Auftrennung der komplexen Mischung aus mRNA-Molekülen in Fraktionen, jeweils bestehend aus 50-100 Genen der totalen Population, wurde erreicht durch: 1) reverse Transkription der mRNA in einzelsträngige cDNA mit 12 sogenannten 3'-Anker-Primern der Form  $T_{11}VN$  (wobei  $T_{11}$  = elf aufeinanderfolgende T's,  $V = A, C, G$ ;  $N = A, C, G, T$ ); 2) PCR-Amplifikation jeder einzelnen cDNA-Fraktion mit dem entsprechenden 3'-Ankerprimer und einem willkürlich gewählten 10 Nukleotide umfassenden 5'-Oligomer in Gegenwart von radioaktiv markierten Desoxyribonukleotiden. Die Produkte wurden auf Sequenziergelen aufgetrennt und es wurden 50-100 Banden im Grössenbereich von 100-500 Nukleotiden beobachtet. Die Banden resultierten aus der Amplifikation von cDNAs, die mit den 3'-Enden von mRNAs korrespondieren, welche das Komplement des 3'-Ankerprimers sowie des willkürlich gewählten 5'-Oligomers enthielten. Für jedes Primerpaar waren die Muster der aus beiden cDNAs amplifizierten Banden ähnlich, wobei die Intensitäten von etwa 80 % der Banden nicht unterscheidbar waren. Bestimmte Banden traten in dem einen oder anderen PCR-

Ansatz stärker in Erscheinung, während einige lediglich in einem der beiden Ansätze nachweisbar waren.

Wenn sämtliche der für Säugetiere erwarteten 50.000-100.000 verschiedenen mRNAs mit den willkürlich gewählten 5'-Primern (arbitrary primer) erfassbar wären, dann wäre eine Anzahl von 80-95 solcher Oligonukleotide und ca. 1.000 PCRs notwendig, um mit hoher Wahrscheinlichkeit etwa zwei Drittel dieser mRNAs zu detektieren. Aus zahlreichen Untersuchungen der letzten Jahre hat sich gezeigt, dass das oben dargestellte Verfahren zu einer hohen Rate (bis 90 %) von falsch-positiven Signalen führt.

Die WO 95/13369 offenbart ein Verfahren (TOGA - Total Gene Expression Analysis) zur gleichzeitigen Identifizierung differentiell exprimierter mRNAs und zur Messung deren relativer Konzentrationen. Das Verfahren basiert auf der Erstellung doppelsträngiger cDNA aus isolierter mRNA unter der Verwendung eines spezifischen Sets an Oligo (dT)-Primern. Dabei wird eine Mischung aus 12 Ankerprimern mit der folgenden Struktur eingesetzt: von 5' beginnend folgt einem "Stuffer"- oder "Heel"-Fragment von 4-40 Basen eine Erkennungssequenz für eine Restriktionsendonuklease (typischerweise NotI), 7-40 dT-Nukleotide und schliesslich zwei "Ankerbasen" V, N am 3'-Ende des Primers. Dabei bezeichnet V ein Desoxyribonukleotid der Gruppe dA, dC oder dG, während N die Desoxyribonukleotide dA, dC, dG und dT definiert. Die auf diese Weise erhaltene cDNA wird anschliessend mit einem Restriktionsenzym, welches 4 Basen als Sequenz für die Schnittstelle erkennt (z.B. MspI), vollständig verdaut, mit NotI geschnitten und in einen entsprechend behandelten Plasmid-Vektor kloniert. Die Orientierung des Inserts ist dabei "antisense" relativ zu einem Vektor-kodierten, Bakteriophagen-spezifischen Promotor (typischerweise T3). Die Ligationen werden in einen E.coli-Stamm transformiert, wodurch cDNA-Banken generiert werden. Die Plasmid-DNA dieser cDNA-Bibliotheken wird isoliert und mit Hilfe von Kombinations-Verdau durch 6 unterschiedliche Restriktionsenzyme, die von den oben verwendeten verschieden sind, linearisiert. Die linearisierte cDNA wird durch die T3-Polymerase in cRNA übersetzt und anschliessend in 16 Sub-Fractionen einzelsträngiger cDNA transkribiert. Dabei wird eine thermostabile reverse Transkriptase bei hoher Temperatur und je einer von 16 verschiedenen cRNA-Primern, dessen beiden 3'-Nukleotide aus einer vollständigen Permutation der 4 möglichen Desoxyribonukleotide dA, dC, dG und dT besteht, verwendet. Die Produkte der 16 cDNA-Fractionen werden als Template für PCR eingesetzt unter Verwendung eines 3'-Oligonukleotides, welches einer

Vektor-Sequenz nahe der Klonierungsstelle des Inserts entspricht, und einem 5'-Oligomer, das einem der 16 cDNA-Synthese-Primer mit zusätzlich zwei 3'-Nukleotiden der vollständigen Permutation der 4 möglichen Desoxyribonukleotide dA, dC, dG und dT entspricht. So werden bis zu 256 verschiedene Pools generiert, deren radioaktiv markierte Banden (<sup>35</sup>S-dATP oder <sup>32</sup>P-dCTP) auf Polyacrylamid-Gelen analysiert werden. Aufgrund der erhaltenen Information über die Länge und Zusammensetzung der 8 identifizierten Basen einer markierten Bande soll es theoretisch möglich sein, ohne Klonierungs- und Sequenzierungsschritte auf die Identifikation des zugehörigen Gens in einer kompletten Datenbank zu schliessen.

Die oben dargelegte Methodik wirft folgende Probleme im Zusammenhang mit der erfindungsgemäss angestrebten hohen Spezifität, Selektivität und Reproduzierbarkeit auf:

- 1) potentieller Verlust von cDNA-Sequenzen durch NotI-Verdau;
- 2) potentieller Verlust von cDNA-Sequenzen durch Vektor-Ligation;
- 3) potentieller Verlust von cDNA-Sequenzen durch Transformation in E.coli und unterschiedliche Amplifikationsraten für verschiedene cDNA-Inserts;
- 4) Kontamination der PCR-Templates mit bakterieller genomischer DNA nach Plasmid-Amplifikation und Reinigung aus E.coli;
- 5) Kontamination der Templates für die T3-RNA-Polymerase-Reaktion mit Insert-freier Plasmid-DNA;
- 6) Verlust von cDNA-Inserts durch kombinierte Linearisierungs-Verdaus mit 6 unterschiedlichen Restriktionsendonukleasen;
- 7) Verlust von cDNA-Sequenzen durch cRNA-Synthese-Amplifikation;
- 8) Verlust von cDNA-Sequenzen durch die zweite cDNA-Synthese-Amplifikation ;
- 9) Fragliche Spezifität der thermostabilen reversen Transkriptase für die permutierten Primer, die für die zweite cDNA-Synthese eingesetzt werden, da die Selektivität der reversen Transkriptase bei Basen-Misspaarung im allgemeinen ca 10-1.000-fach (für AMV-RT) unterhalb der Selektivität der Taq-Polymerase liegt (L.V. Mendelman et al. 1990, J. Biol. Chem. 265, 2338-2346);
- 10) Fragliche Selektivität der Taq-Polymerase für die verwendeten permutierten 5'-Oligomere in der PCR unter den beschriebenen Bedingungen, da Basenmisspaarungen bei den eingesetzten Primern toleriert werden. Die korrekte Analyse der Informationen ist deshalb nicht gewährleistet.

M. Matz et al. beschreiben ein weiteres Verfahren (ODD - ordered differential display - 1997, Nucl. Acid. Res. 25, 2541-2542) zur Identifizierung differentiell exprimierter Gene, welches auf der PCR-Amplifikation durch Adaptor-spezifische Oligonukleotide und dem Effekt der PCR-Suppression beruht. Hierbei wird doppelsträngige cDNA unter Verwendung eines Oligonukleotides der Struktur "Heel-dT(13)" erzeugt, wobei "Heel" eine Folge von 12 Basen darstellt. Die cDNA wird von einem Restriktionsenzym mit einer 4 Basen umfassenden Erkennungssequenz (RsaI) vollständig verdaut und mit einem "Pseudo-Doppelstrang-Adaptor" ligiert. Dieses Molekül besteht aus einem längeren (39 Basen) und einem kürzeren (12 Basen komplementär zum 3'-Ende des längeren) Oligomer, welche unter geeigneten Bedingungen gegeneinander hybridisiert werden. Die 5'-Enden der Oligonukleotide sind dabei nicht phosphoryliert. Theoretisch ist so in einer 1. PCR die spezifische Amplifikation der 3'-Enden der cDNA mit Hilfe des cDNA-Synthese-Primers und eines Primers, der dem 5'-Ende des längeren Adaptor-Oligomers entspricht, bei hoher Annealing-Temperatur in der PCR (65°C) möglich. Eine Auftrennung der komplexen cDNA in unterschiedliche Fraktionen wird in einer 2. PCR erreicht. Dabei wird der cDNA-Synthese-Primer mit zusätzlich zwei 3'-Nukleotiden der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT und ein Primer, welcher dem 3'-Ende des längeren Adaptor-Oligomers mit zusätzlich zwei 3'-Nukleotiden der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT entspricht, verwendet. Für eine erhöhte Spezifität des Adaptor-Primers wurde eine artifizielle Misspaarung (Mismatch) in das Oligonukleotid an der Position -4 relativ zum 3'-Ende des Primers eingeführt.

Die oben dargelegte Methodik wirft folgende Probleme im Zusammenhang mit der erfindungsgemäss angestrebten hohen Spezifität, Selektivität und Reproduzierbarkeit auf:

- 1) Die Verwendung eines nicht-geankerten cDNA-Synthese-Primers erlaubt keine Garantie der Reproduzierbarkeit der gefundenen Fragmentlängen;
- 2) Die Amplifikation der 3'-Enden der cDNA unter Verwendung eines Oligomers der Struktur "Heel-dT(13)" bei hohen Annealing-Temperaturen in der PCR ist nicht reproduzierbar;
- 3) Die selektive Amplifikation der 3'-Enden der cDNA durch Verwendung eines "Pseudo-Doppelstrang-Adaptors" ist nicht gesichert;
- 4) Die erhöhte Spezifität in der 2. PCR für die permutierten 5'-Oligonukleotide durch die Einführung eines artifiziellen Mismatches ist unbefriedigend;

- 5) Die in der 2. PCR eingesetzten 3'-Oligomere enthalten keinen artifiziellen Mismatch und sind daher nicht ausreichend selektiv für die Primer-Permutation;
- 6) Durch die spezielle Anordnung der PCR ist die aus der Gel-Analyse erhaltene Information über die einzelnen DNA-Fragmente (Fragmentlänge und 6 bekannte Nukleotide) ohne zusätzliche Klonierungs- bzw. Sequenzierungsschritte nicht ausreichend.

Y. Prashar und S. Weissman (1996, Proc. Natl. Acad. Sci. USA 93, 659-663) beschreiben ein Verfahren, bei welchem doppelsträngige cDNA mittels 12 Oligonukleotiden hergestellt wird, die folgende Struktur aufweisen: von 5' beginnend folgt einer "Heel"-Struktur eine Folge von 18 dT-Nukleotiden und zwei "Ankerbasen" V, N am 3'-Ende des Primers. Dabei bezeichnet V ein Desoxyribonukleotid der Gruppe dA, dC oder dG, während N die Desoxyribonukleotide dA, dC, dG und dT definiert. Die cDNA-Synthese erfolgt bei einer Temperatur von 50°C und soll eine Aufteilung der komplexen Mischung in 12 unterschiedliche Pools ermöglichen. Nach vollständigem Verdau der cDNA mit verschiedenen Restriktionsendonukleasen, welche sechs Nukleotide als Sequenz für die Schnittstelle erkennen, werden die entstandenen DNA-Fragmente mit einem Adaptor versehen, welcher die Struktur eines Ypsilons aufweist. In der nachfolgenden PCR wird ein 3'-Primer verwendet, der an die "Heel"-Struktur der cDNA bindet. Als 5'-Primer wird ein Oligonukleotid eingesetzt, dessen Bindungsstelle in der äusseren Region des Ypsilon-Adaptors liegt und die erst entsteht, wenn in einer ersten Synthese der komplementäre Strang zu dieser Region gebildet wird. Alle Fragmente, welche auf beiden Seiten einen Ypsilon-Adaptor besitzen, können nicht amplifiziert werden.

Die von den Autoren behauptete Auftrennung der cDNA-Produkte in unterschiedliche Fraktionen ist durch die verwendeten, am 3'-Ende permutierten cDNA-Synthese-Primer nicht reproduzierbar. Zudem können durch die Verwendung der speziellen Adaptor-Struktur keine permutierten PCR-5'-Primer eingesetzt werden.

Die WO 97/29211 beschreibt die Technik 'Restriction Display (RD-PCR)', bei welcher doppelsträngige cDNA mittels 12 Oligonukleotiden hergestellt wird. Diese Primer weisen folgende Struktur auf: von 5' beginnend folgen einer "Heel"-Struktur zwei Desoxynukleotide der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT, eine Folge von 17 dT-Nukleotiden und zwei "Ankerbasen" V, N am 3'-Ende des Primers. Dabei bezeichnet V ein Desoxyribonukleotid der Gruppe dA, dC oder dG, während N die



Desoxyribonukleotide dA, dC, dG und dT definiert. Nach vollständigem Verdau der cDNA mit einem oder mehreren Restriktions-Endonukleasen wird ein Adaptor-Molekül an die cDNA-Fragmente ligiert. In einer sich anschliessenden PCR wird ein 3'-Primer verwendet, der selektiv an die "Heel"-Struktur der cDNA bindet und zusätzlich zwei 3'-Nukleotide V, N am 3'-Ende des Primers besitzt. Dabei bezeichnet V ein Desoxyribonukleotid der Gruppe dA, dC oder dG, während N die Desoxyribonukleotide dA, dC, dG und dT definiert. Als 5'-Primer wird ein der 3'-Sequenz des Adaptor-Primers entsprechendes Oligonukleotid eingesetzt, welches zusätzlich ein 3'-Nukleotid oder zwei 3'-Nukleotide oder drei 3'-Nukleotide der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT aufweist. Die PCR wird dabei mit unterschiedlichen Permutations-Kombinationen derart durchgeführt, dass in einer ersten PCR (oder in den ersten 10-25 PCR-Zyklen) 5'-Primer mit nur einer Permutation und anschliessend in einer zweiten PCR (oder in den restlichen PCR-Zyklen) 5'-Primer mit nur zwei oder drei Permutationen eingesetzt werden. Hierdurch soll die Selektivität für die verschiedenen 5'-Primer-Permutationen deutlich erhöht werden.

Die oben dargelegte Methodik wirft folgende Probleme im Zusammenhang mit der erfindungsgemäss angestrebten hohen Spezifität, Selektivität und Reproduzierbarkeit auf:

- 1) Die Verwendung von verschiedenen cDNA-Synthese-Primern mit 3'-Permutationen ist nicht selektiv und erlaubt daher keine Garantie der Reproduzierbarkeit der Methode;
- 2) Die Aufteilung des Amplifikationsschrittes in mehrere PCR-Runden mit 5'-Primern, welche eine unterschiedliche Anzahl an 3'-Permutationen aufweisen, ist allein für sich nicht Sequenz-selektiv genug, um die Technik z.B. in einer Datenbank-orientierten Gen-Expressions-Analyse einzusetzen.

Kato beschreibt ein Verfahren ('molecular indexing' - 1995, Nucl. Acids Res., Vol. 23, 3685-90 und 1996, Nucl. Acids Res., Vol. 24, 394-95), das auf dem Verdau der doppelsträngigen cDNA mit Restriktionsendonukleasen der Klasse IIS beruht. Diese generieren 5'-Überhänge der cDNA von unbekannter Sequenz. Es werden dann 64 biotinylierte Adaptoren, deren Nukleotide 2-4 (relativ zum 5'-Ende) ihrer 5'-Überhänge komplementär zu je einem 64-stel des gesamten cDNA-Pools sind, mit DNA-Ligase aus E.coli ligiert. Das jeweilige 5'-Nukleotid der Adaptoren-Überhänge bleibt undefiniert. Die dabei ligierten cDNA-Fragmente werden über die Bindung an Streptavidin-gekoppelte magnetische Partikel gereinigt. In einer sich anschliessenden PCR werden die Adaptor-ligierten 3'-Enden der cDNA unter Verwendung eines Adaptor-Oligonukleotids und eines Oligo-(dT)-Oligomers, welches am 3'-

Ende um eines der drei Nukleotide dA, dC oder dG verlängert ist, bei geringer Annealing-Temperatur amplifiziert. Die cDNA wird derart in 192 unterschiedliche Pools aufgetrennt. Auch diese Methodik wirft folgende Probleme im Zusammenhang mit der erfindungsgemäss angestrebten hohen Spezifität, Selektivität und Reproduzierbarkeit auf:

- 1) Die Ligation von Adaptoren mit vier Nukleotiden als 5'-Überhang ist ohne weitere Nachbehandlung nicht Permutations-spezifisch genug, um mit Sicherheit die ersten drei oder vier Basen des cDNA-Inserts bestimmen zu können;
- 2) Um eine möglichst hohe Ligationsspezifität zu erreichen, wird nur eine sehr geringe Menge an Adaptor-Molekül in die Ligation eingesetzt, was jedoch zu deutlich verringerter Ligations-Effizienz führt. Dies wiederum hat zur Folge, dass die Sensitivität der gesamten Methode reduziert wird;
- 3) Die Aufteilung der cDNA in Pools unter Verwendung von geankerten Oligo-(dT)-Primern bei geringer Annealing-Temperatur in der PCR ist nicht erfolgreich durchführbar.

Aufgrund der geschilderten Nachteile der oben dargestellten Techniken des Standes der Technik, insbesondere in Bezug auf die mangelnde oder nur geringe Reproduzierbarkeit und Sequenz-Spezifität während der PCR-Amplifikation der 3'-Enden der cDNA, ergibt sich die Notwendigkeit der Entwicklung und Etablierung eines Verfahrens, welches den bekannten Verfahren bezüglich Spezifität, Selektivität, Sensitivität und verlässlicher Reproduzierbarkeit sowie reduzierter Fehlerrate der erhaltenen Ergebnisse deutlich überlegen ist. Diese Kriterien sind insbesondere dann erforderlich, wenn z.B. eine Datenbank-gestützte Analyse zur differentiellen Gen-Expression mittels digitalem Display durchgeführt werden soll.

Erfindungsgemäss wird daher ein Verfahren zur Identifikation und Charakterisierung von mRNA-Molekülen bereitgestellt, welches die folgenden Schritte umfasst:

- (a) Isolierung und Reinigung von PolyA<sup>+</sup>-RNA aus Gewebeproben;
- (b) Synthese von doppelsträngiger cDNA aus den mRNA-Molekülen;
- (c) Verkürzen der cDNA durch enzymatischen Verdau mit Restriktionsendonukleasen;
- (d) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA; und entweder, in einer 1. Alternative,

(e) Auffüllen der 5'-Überhänge der cDNA mit Desoxyribonukleotiden und Klenow-DNA-Polymerase;

(f) selektive Aufreinigung der 3'-Enden der cDNA;

(g) Abtrennung der 3'-ständigen Poly-dA-Nukleotide der cDNA durch enzymatischen Verdau mit einer Restriktionsendonuklease;

(h) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA;

(i) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction);

(j) Auftrennung der Amplifikationsprodukte nach ihrer Länge;

(k) Analyse der Amplifikationsprodukte;

oder, in einer 2. Alternative,

(e) selektive Aufreinigung der 3'-Enden der cDNA;

(f) Abtrennung der 3'-ständigen Poly-dA-Nukleotide der cDNA durch enzymatischen Verdau mit einer Restriktionsendonuklease;

(g) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA;

(h) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction), wobei 5'-Primer verwendet werden, die zwei oder drei Permutationen am 3'-Ende aufweisen, und die eine oder zwei artifizielle Fehlpaarungen im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang enthalten;

(i) Auftrennung der Amplifikationsprodukte nach ihrer Länge;

(j) Analyse der Amplifikationsprodukte;

oder, in einer dritten Alternative,

(e) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction), wobei 5'-Primer verwendet werden, die zwei oder drei Permutationen am 3'-Ende aufweisen, und die eine oder zwei artifizielle Fehlpaarungen im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang enthalten;

(f) Auftrennung der Amplifikationsprodukte nach ihrer Länge;

(g) Analyse der Amplifikationsprodukte;

wobei für die 1. und 2. Alternative in Schritt (b) die Synthese der cDNA-Erststrangmoleküle durch reverse Transkription unter Verwendung eines geankerten Oligo-dT-Nukleotids erfolgt, welches eine 3'-Extension von 2 Basen aufweist, wobei die erste Base dA, dC oder dG, und die zweite Base dA, dC, dG oder dT ist, und welches eine 5'-Extension von 5-15, bevorzugt von 6-15, Basen aufweist, die für die

Schnittstelle einer Restriktionsendonuklease mit der Spalt-Charakteristik 16/14 'downstream' der Erkennungssequenz kodiert; oder wobei für die 3. Alternative in Schritt (b) die Synthese der cDNA-Erststrangmoleküle durch reverse Transkription unter Verwendung eines geankerten Oligo-dT-Nukleotids erfolgt, welches eine 3'-Extension von 2 Basen aufweist, wobei die erste Base dA, dC oder dG, und die zweite Base dA, dC, dG oder dT ist, und welches eine 5'-Extension von 5-15 Basen einer beliebigen Sequenz aufweist.

Enzyme mit der Spaltcharakteristik 16/14 "downstream" der Erkennungssequenz sind dem Fachmann bekannt. Beispiel sind Eco571 und Bsg1 (siehe beispielsweise A. Janulaitis et al., Nucleic Acids Res. 20 (1992), S. 6043-6049; Petrusyte et al., Gene 74 (1988), S. 89-91).

Bevorzugt ist in dem erfindungsgemässen Verfahren das Oligo-dT-Nukleotid vollständig durch 2'-O-methylierte Ribonukleotide substituiert. Alternativ kann das Oligo-dT-Nukleotid aus Standard-Desoxyribonukleotiden bestehen.

Weiter bevorzugt ist in dem erfindungsgemässen Verfahren, dass das Oligo-dT-Nukleotid an seinem freien 5'-Ende und/oder an internen dT-Nukleotiden mit einem Biotinrest über einen C9-Spacer versehen wird.

Eine bevorzugte Ausführungsform betrifft ein erfindungsgemässes Verfahren, wobei in Schritt (h) der 2. Alternative oder in Schritt (e) der 3. Alternative sich die Fehlpaarungen an den Positionen -3, oder -3 und -4, oder -4 und -5 im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang befinden.

Eine weitere bevorzugte Ausführungsform des erfindungsgemässen Verfahrens ist dadurch gekennzeichnet, dass in Schritt (c) ein Restriktionsenzym der Klasse IIS verwendet wird, welches 5 Nukleotide als Erkennungssequenz besitzt und einen aus 2 - 4, insbesondere 4, Nukleotiden, die nicht Teil der Erkennungssequenz sind, bestehenden Überhang der geschnittenen cDNA-Fragmente generiert.

Eine weitere bevorzugte Ausführungsform des erfindungsgemässen Verfahrens ist dadurch gekennzeichnet, dass in Schritt (f) der 1. Alternative oder in Schritt (e) der 2. Alternative die

selektive Aufreinigung der 3'-Enden der cDNA unter Verwendung von paramagnetischen 'beads' erfolgt, die das Biotin-bindende Molekül Streptavidin gekoppelt haben.

Bevorzugt wird in Schritt (g) der 1. Alternative oder in Schritt (f) der 2. Alternative des erfindungsgemässen Verfahrens ein Restriktionsenzym der Klasse IIS verwendet wird.

Weiter bevorzugt werden die Produkte aus Schritt (d) oder aus Schritt (h) der 1. Alternative oder aus Schritt (g) der 2. Alternative des erfindungsgemässen Verfahrens vor der Amplifizierung gemäss Schritt (i) mit einer Nuklease, ausgewählt aus der Gruppe bestehend aus T4-Endonuclease VII, S1-Nuclease, und Mung Bean Nuclease, inkubiert werden.

Ferner ist das erfindungsgemässe Verfahren bevorzugterweise dadurch gekennzeichnet, dass in Schritt (i) der ersten Alternative oder in Schritt (h) der 2. Alternative oder in Schritt (e) der 3. Alternative die Amplifizierung der cDNA-Fragmente unter Verwendung von Oligonukleotiden erfolgt, die an den komplementären Strang des 'sense'-Oligomers der ligierten Adaptor-Moleküle hybridisieren.

Das erfindungsgemässe Verfahren ist weiter bevorzugt dadurch gekennzeichnet, dass in Schritt (k) der ersten Alternative oder in Schritt (j) der 2. Alternative oder in Schritt (g) der 3. Alternative die Analyse anhand der unterschiedlichen Längen der Produkte sowie in Kenntnis der durch die Manipulation bekannten Basensequenz von 9 oder 10 Nukleotiden erfolgt.

Eine weitere bevorzugte Ausführungsform betrifft die Anwendung des erfindungsgemässen Verfahrens nach einem der vorhergehenden Ansprüche zur gegebenenfalls rechnergestützten Identifizierung und Isolierung sowie Analyse neuer Gene.

Das erfindungsgemässe Verfahren DEPD wurde entwickelt, um die Fehlerrate bei der Identifizierung einiger Nukleotide der cDNA-Fragmente, welche wiederum die Identifizierung des kodierten Gens in einer geeigneten Datenbank zulassen, deutlich zu reduzieren. Die verbesserte Leistungsfähigkeit des erfindungsgemässen Verfahrens ergibt sich aus der Anwendung spezifischer Ligations-Techniken in geeigneter Kombination mit permutations-spezifischer Mismatch-PCR. Dabei wird nach selektiver Aufreinigung der 3'-Enden der cDNAs an das 5'- und 3'-Ende der Fragmente ein Adaptor-Molekül ligiert. In der folgenden

PCR werden Primer eingesetzt, die vorzugsweise je zwei Basen als Permutation aufweisen. Die Permutations-Spezifität der PCR-Primer kann bei hoher Annealing-Temperatur und vorzugsweise bei gleichzeitiger Einführung von artifiziellen Template-Fehlpaarungen an selektive Oligonukleotid-Stellen deutlich erhöht werden.

Durch die erfindungsgemässe Kombination mehrerer Einzelschritte, welche zur Bestimmung von Nukleotid-Sequenzen mittels PCR eingesetzt werden, zeigt sich ein Synergie-Effekt was die Spezifität der Technik betrifft. Es ergibt sich eine Art Kontroll-Mechanismus für die einzelnen Prozeduren, da jeder Fehler, der z.B. in der Ligation erfolgt, durch die selektive PCR mittels Mismatch-Primern korrigiert werden kann. Die Permutations-Selektivität der PCR wurde durch die bevorzugte Einführung von definierten Template-Fehlpaarungen deutlich gesteigert.

Nach einer weiteren bevorzugten Ausführungsform erhält man durch die Verwendung von Permutations-Primern auf beiden Seiten des cDNA-Templates einen Fehler-korrigierenden Effekt, da solche Amplifikations-Produkte, welche fälschlicherweise mittels eines Oligomers in einer der ersten PCR-Runden erstellt wurden, in ihrer weiteren Vermehrung supprimiert werden können, falls der Gegen-Primer nur die korrekte Permutation amplifiziert. Ferner ist es gemäss einer weiteren bevorzugten Ausführungsform durch den geeigneten, kombinierten Einsatz von Ligations- und PCR-Permutations-Technik bei der DEPD-Methode möglich, 9 oder 10 Nukleotide und die Länge eines amplifizierten Fragmentes zu definieren. Diese Information sollte zur sicheren Identifikation eines Gens, vorzugsweise mittels Datenbank-Analyse, auch dann hinreichend sein, falls sich ein Fehler bei der Bestimmung der Basen-Sequenz des cDNA-Fragmentes ergeben sollte.

Die Anwendung des erfindungsgemässen DEPD-Verfahrens ermöglicht u.a. eine umfassende Analyse des Zusammenspiels aller in einem definierten System und/oder in einer definierten Situation involvierten Gene auf der Ebene der mRNA-Expressionsmuster, wobei für spezifische und reproduzierbare Ergebnisse lediglich geringe Mengen an Gewebe oder Zellen benötigt werden. Das Verfahren kann in einer Vielzahl von Anwendungsbereichen eingesetzt werden. Hierzu zählen z.B. der Vergleich von Organen, Geweben, Gewebeteilen, oder von erkrankten Geweben oder Gewebeteilen mit entsprechendem gesunden Material, ggfs. auch im Rahmen einer vergleichenden Untersuchung unter Verwendung von pharmazeutischen Wirkstoffen gegenüber entsprechenden Kontrollen

ohne Wirkstoffapplikation. Ferner ermöglicht das erfindungsgemässe Verfahren einen Vergleich von definierten Zuständen in Tiermodellen, wobei hier insbesondere vergleichende Analysen von Organen, Geweben, Gewebeteilen oder erkrankten Geweben bzw. Gewebeteilen gegenüber entsprechendem gesunden Material bevorzugt sind. Weitere Anwendungsgebiete betreffen die Analyse transgener Tiere, welche auch sog. 'knock-out'-Tiere einschliessen, sowie die phänotypische Evaluierung des Einsatzes von Antikörpern, Antisense- und Ribozym-Oligonukleotiden und vergleichbarer Mittel, die im Rahmen von funktionellen Ansätzen zur Aufklärung der Relevanz bestimmter Gene eingesetzt werden.

Der Einsatz einer solchen Technik erlaubt es, die Screening-Geschwindigkeit von zu untersuchendem Material deutlich zu erhöhen, da eine Isolierung differentiell exprimierter Gene mit anschliessender Klonierung und Sequenzierung nicht mehr notwendig ist. Es ist daher möglich, eine viel grössere Zahl an Proben, bzw. weitaus mehr verschiedene Stadien einer Probe (z.B. zeitliche Verläufe von Veränderungen der Genexpression über mehrere Stunden/Tage etc.) zu untersuchen. Dies ist nicht nur im Hinblick auf die Entdeckung neuer möglicher 'drug targets' wichtig, sondern auch besonders unter dem Aspekt der Aufklärung von Wirkmechanismen potentieller therapeutischer Substanzen, da hier bei detaillierter Untersuchung besonders viel Probenmaterial anfallen kann.

Ein weiterer Gesichtspunkt der vorliegenden Erfindung ist deren bevorzugte Anwendung im Rahmen eines Datenbank-orientierten Gen-Expressions-Analyse-Verfahrens. Im Gegensatz zum enormen technischen Aufwand bereits etablierter Verfahren mit ähnlich hoher Leistungsfähigkeit wie z.B. der Chip-Hybridisierungs-Technologie (z.B. Affymetrix, USA), von der Chip-Herstellung über die Auswertungstechnik der Chip-Hybridisierung bis zur automatisierten Datenanalyse, stellt das erfindungsgemässe Verfahren eine überraschend einfache und kostengünstige Alternative dar, welche bekannten Verfahren des eingangs geschilderten Typs qualitativ deutlich überlegen ist.

Nachfolgend wird das erfindungsgemässe Verfahren sowie bevorzugte Ausführungsformen desselben im Detail beschrieben. Hinsichtlich detaillierter Angaben zu etablierten Standard-Verfahren wird z.B. verwiesen auf J. Sambrook et al. 1989: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

## Beispiele

### 1) mRNA-Isolation

Die Isolierung und Aufreinigung von Gesamt-RNA aus zu untersuchenden Geweben, Gewebeteilen, Biopsieproben, Zellen etc. erfolgt nach beschriebenen Standard-Verfahren. Zur vollständigen Abtrennung von eventuellen Kontaminationen der isolierten RNA mit genomischer DNA wird ein enzymatischer Verdau mit DNaseI (Boehringer Mannheim, Deutschland) durchgeführt. Es folgt dann die Reinigung von PolyA<sup>+</sup>-mRNA aus der totalen RNA mittels Oligo(dT)-gekoppelter magnetischer Partikel (Oligo(dT) magnetic beads, Promega, WI, USA).

### 2) cDNA-Synthese

Doppelsträngige cDNA wird aus mRNA synthetisiert unter Verwendung einer Mischung aus 12 Anker-Primern mit folgender Struktur: von 5' beginnend folgt auf ein "Heel"-Fragment (5'-Extension) aus 5-15 Basen, beispielsweise 5 oder 6 Basen, eine Erkennungssequenz für eines der Restriktions-Enzyme BsgI oder Eco57I. Anschliessend folgt eine Abfolge von 14 dT-Nukleotiden und die zwei Anker-Nukleotide V, N am 3'-Ende des Primers. Dabei bezeichnet V ein Desoxyribonukleotid der Gruppe dA, dC oder dG, während N die Desoxyribonukleotide dA, dC, dG und dT definiert. Die Desoxyribonukleotide der dT-Folge, gegebenenfalls auch des "Heel"-Fragmentes, sind hierbei komplett substituiert durch 2'-O-methylierte Ribonukleotide. Der cDNA-Synthese-Primer ist am 5'-Ende und/oder an einem oder mehreren internen dT-Nukleotiden mit einem Biotin-Rest (C9-Spacer) versehen.

Alternativ kann die cDNA-Synthese, statt mit 2'-O-methylierten Ribonukleotiden, mit einer Mischung aus 12 nicht-modifizierten, d.h. aus Desoxyribonukleotiden bestehenden, Anker-Primern durchgeführt werden.

### 3) Erster Verdau der cDNA mit Restriktionsendonukleasen

Vollständiger Verdau der cDNA mit einem Restriktionsenzym des Typs IIS, welches 5 Basen als Sequenz für die Schnittstelle erkennt (z.B.: FokI, Bsm AI, Bsm FI oder Bbv I) und zwei bis vier unbekannte Nukleotide als Überhang der cDNA generiert. In den weiteren Ausführungen werden beispielhaft Enzyme angeführt, die einen Nukleotid-Überhang von 4 Basen generieren. Bei kürzeren Überhängen müssen die zu ligierenden Adaptoren entsprechend verkürzt werden.



#### 4) Erste Adaptoren-Ligation

a) Durch den Enzym-Verdau entsteht ein 5'-Überhang von vier unbekannten Nukleotiden in der cDNA. Sechzehn Adaptor-Moleküle, bestehend aus zwei Oligomeren (nicht 5'-phosphoryliert, "Pseudo-Doppelstrang"-Adaptoren) werden nach erfolgtem Restriktions-Verdau mit der cDNA ligiert. Typischerweise besteht der Adaptor aus zwei unterschiedlich langen Oligonukleotiden, wobei das längere aus 25-35, das kürzere (komplementär zum 3'-Ende des längeren) aus 8-25 Basen, insbesondere 8-12 Basen, besteht. Der Adaptor entsteht durch Hybridisierung der beiden Oligonukleotide gegeneinander unter geeigneten Bedingungen und bildet einen 4-Nukleotid-Überhang, wobei das Nukleotid 3 bzw. 4 ('innere' Nukleotide - relativ zum 5'-Ende des antisense-Oligomers) eines der vier möglichen Desoxyribonukleotide dA, dC, dG und dT sein kann, während die 'äusseren' Nukleotide 1 bzw. 2 (relativ zum 5'-Ende des antisense-Oligomers) undefiniert bleiben, d.h. immer als Mischung aus der vollständigen Kombination aller vier Desoxyribonukleotide bestehen. So entstehen 16 permutierte Adaptoren, deren 'innere' beiden Nukleotide die Spezifität der Ligation bestimmen. Zur Sicherstellung der korrekten Adaptoren-Ligation werden die Ligations-Ansätze nachträglich mit einer Nuklease, ausgewählt aus der Gruppe bestehend aus T4-Endonuclease VII, S1-Nuclease, und Mung Bean Nuklease, inkubiert.

Die Ligation der 16 permutierten Adaptor-Moleküle erfolgt bei 4-37°C, 1-16 Stunden, 0-150 mM Na-Acetat, 1-5 Einheiten T4-DNA-Ligase (oder Taq-DNA-Ligase oder E.coli-DNA-Ligase) in geeignetem Puffer. Jeder der 16 Ligationsansätze wird nach Aufreinigung in einer PCR als Template eingesetzt (siehe (8a)).

b) Alternativ werden statt 16 schliesslich 64 verschiedene "Pseudo-Doppelstrang"-Adaptoren, welche einen 4-Nukleotid-Überhang bilden, mit der cDNA ligiert. Dabei setzen sich die 64 Adaptoren-Überhänge aus allen Desoxyribonukleotid-Kombinationen der Basen 2-4 (relativ zum 5'-Ende des antisense-Oligomers) zusammen, während das 'äussere' Nukleotid (relativ zum 5'-Ende des antisense-Oligomers) als Mischung aller vier Nukleotide eingesetzt wird. Damit bestimmen die 'inneren' drei Nukleotide die Spezifität der Ligation. Dieses Verfahren beruht auf der Erkenntnis, dass die DNA-Ligase aus E.coli die drei ersten Basen des Adaptor-Überhangs in der Ligation zu diskriminieren vermag. Diese erfolgt unter den bereits oben beschriebenen Bedingungen. Zur Sicherstellung der korrekten Adaptoren-Ligation werden die Ligations-Ansätze nachträglich mit einer Nuklease, ausgewählt aus der

Gruppe bestehend aus T4-Endonuclease VII, S1-Nuclease, und Mung Bean Nuklease, inkubiert.

c) Alternativ werden statt der Ligation von 16 oder 64 verschiedenen "Pseudo-Doppelstrang"-Adaptoren die 5'-Überhänge der cDNA sukzessive mit Desoxyribonukleotiden unter Verwendung von Klenow-DNA-Polymerase aufgefüllt. Dies ist möglich durch den Einsatz von Didesoxyribonukleotiden (ddNTP) und kompetitiven Adaptor-Molekülen. Zudem werden die 3'-Enden der cDNA nicht von den magnetischen Partikeln eluiert, um die eingesetzten Nukleotide nach jedem Schritt entfernen zu können. Soll beispielsweise der Basen-Überhang 5'-GGTT-3' aufgefüllt werden, wird die cDNA zuerst mit ddC-Nukleotiden inkubiert, um alle Überhänge, welche mit einem dG beginnen zu blockieren. Nach Entfernung der Didesoxyribonukleotide wird die cDNA mit dA-Nukleotiden inkubiert. Danach erfolgt die Ligation eines Adaptor-Moleküls, das ein 'blunt' und ein 'sticky'-Ende besitzt, an diejenigen cNDAs, die bis dahin einen aufgefüllten 5'-Überhang aufweisen. Im gleichen Schritt werden zwei kompetitive Adaptor-Moleküle ligiert. Diese weisen ebenfalls ein 'sticky'-Ende auf und einen 5'-Überhang, der aus einem dC- bzw. drei dC-Molekülen besteht. Diese blockieren die cDNA-5'-Überhänge 5'-GTTT-3' bzw. 5'-GGGT-3'. Es folgt eine weitere Aufreinigung der Streptavidin-gebundenen cDNA-Fragmente, die dann mit dC-Nukleotiden inkubiert wird. Anschliessend erfolgt die Ligation eines Adaptor-Moleküls mit einem 'blunt'-Ende, welcher nur an die vollständig und korrekt aufgefüllte cDNA ligieren kann. Auf die Art und Weise wird jeder der 256 möglichen cDNA-5'-Überhänge mit DNA-Polymerase komplett aufgefüllt und anschliessend in eine PCR eingesetzt.

##### 5) Selektive Aufreinigung der 3'-Enden der cDNA

Zur Gewährleistung der selektiven Amplifikation der 3'-Enden der cDNA wird die spezifische Aufreinigung dieser Fragmente aus dem Gesamt-Pool der komplexen cDNA durchgeführt. Die 3'-Enden der cDNA können vorzugsweise mit Hilfe von magnetischen Partikeln (magnetic beads), welche an Streptavidin gekoppelt sind, aus dem cDNA-Gemisch selektiv aufgereinigt werden (s. auch Biomagnetic Techniques in Molecular Biology, Dynal, N-0212 Oslo, Norwegen). Diese Art der Aufreinigung stellt sicher, dass in der nachfolgenden PCR keine unspezifische Amplifikation von internen (d.h. Nicht-3'-Enden) cDNA-Fragmenten erfolgt. Die Elution der cDNA von den magnetischen Partikeln erfolgt entweder durch die Extraktion mit organischen Lösungsmitteln bei hoher Temperatur

(65°C) oder durch enzymatischen Verdau mit einer der Restriktionsendonukleasen BsgI oder Eco57I (s.u.).

6) Zweiter Verdau der cDNA mit Restriktionsendonukleasen für die Alternative (4b)

Vollständiger Verdau der ligierten cDNA mit einem der Restriktionsenzyme BsgI oder Eco57I. Dabei werden die Oligo-dT-Nukleotide des 3'-Endes der cDNA komplett abgespalten und ein 2-Nukleotid-Überhang der letzten beiden 3'-Basen (V, N) der cDNA, welche 5' relativ zum Poly-dA-Stretch der mRNA liegen, generiert.

7) Zweite Adaptoren-Ligation für die Alternative (4b)

Je vier Adaptoren-Moleküle, bestehend aus zwei Oligomeren, werden nach erfolgreichem Restriktions-Verdau mit je einem Ligations-Ansatz aus (4b) ligiert. Typischerweise besteht der Adaptor aus zwei unterschiedlich langen Oligonukleotiden, wobei das längere aus 25-35, das kürzere (komplementär zum 3'-Ende des längeren, 5'-phosphoryliert) aus 22-30 Basen besteht. Der Adaptor entsteht durch Hybridisierung der beiden Oligonukleotide gegeneinander unter geeigneten Bedingungen und bildet ein 2-Nukleotid-Überhang, wobei das 'äussere' 3'-Nukleotid eines der vier Desoxyribonukleotide dA, dC, dG oder dT sein kann, während das 'innere' Nukleotid aus einer Mischung der drei Basen dA, dC oder dG definiert ist. Somit bestimmt das 'äussere' 3'-Nukleotid des Adaptor-Überhangs die Spezifität der Ligation. Die Ligations-Bedingungen sind wie oben beschrieben. Jeder der 256 Ligationsansätze wird nach Aufreinigung in eine PCR als Template eingesetzt (siehe (8b)).

8a) PCR für die Alternative (4a)

In insgesamt 256 - 1024 PCRs wird als 3'-Primer der komplett durch 2'-O-methylierte Ribonukleotide substituierte cDNA-Synthese-Primer (Mischung aus 12 Anker-Primern) verwendet. Dabei kann durch die erhöhte Dissoziations-Temperatur der modifizierten Basen des Oligonukleotides eine deutlich höhere (bis zu 40 %, s. z.B. L.L. Cummins 1995, Nucl. Acids Res., 23, 2019-2024) Annealing-Temperatur in der PCR verwendet werden als mit einem nicht-substituierten Primer.

Ein Aliquot der unter Punkt (4a) generierten Ligations-Ansätze wird in der PCR eingesetzt. Als 5'-Primer werden entweder in einer ersten PCR-Runde 16 - 64 Oligomere der Länge 18-27 Basen, die dem 3'-Ende des Sense-Adaptor-Oligonukleotids mit zusätzlich zwei oder

drei 3'-Nukleotiden der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT entsprechen, verwendet. Die hier eingesetzten Primer-Permutationen entsprechen dabei den in den Ligations-Ansätzen unter (4a) definierten Basen 3 - 4 oder 2 - 4 der Adaptoren-Überhänge. In einer zweiten sich anschliessenden PCR-Runde werden pro erste PCR weitere 16 oder 64 PCRs durchgeführt mit demselben 3'-Primer und je 16 5'-Oligonukleotiden der Länge 18-27 Basen, die dem 5'-Primer der ersten PCR entsprechen, jedoch um zwei oder drei 5'-Nukleotide der vollständigen Permutation aller vier Desoxyribonukleotide dA, dC, dG und dT verlängert sind.

Zur Gewährleistung der Spezifität der Amplifikation der verschiedenen Permutationen in der PCR werden selektive PCR-Primer verwendet, welche mehrere artifiziell eingeführte Fehlpaarungen gegenüber der Template-DNA enthalten können. Diese Mismatches können sich an irgendeiner Stelle des Oligomers befinden, wobei 2 Mismatches an den Positionen - 2 und -3 oder 1 Mismatch an der Position -1 relativ zum 3'-Ende des Primers (Position 0) bevorzugt sind.

Das PCR-Profil ist typischerweise: 3 min, 95°C gefolgt von 20-40 Zyklen mit 45 sec, 95°C, 45 sec, 65°C, 60 sec, 72°C und einer abschliessenden Extension für 60 sec bei 72°C, und wird mit Hilfe von radioaktiv oder fluoreszierend markierten PCR-Primern durchgeführt.

#### 8b) Weitere Möglichkeit einer PCR für die Alternative (4a)

Alternativ wird die PCR wie unter 8a) beschrieben durchgeführt, wobei jedoch ein nicht-modifizierter cDNA-Synthese-Primer (Mischung aus 12 Anker-Primern), welcher aus Standard-Desoxyribonukleotiden besteht, eingesetzt.

#### 8c) PCR für die Alternative (4b)

In insgesamt 256 PCRs werden 64 5'-Primer der Länge 18-27 Basen, die dem 3'-Ende des Sense-Oligonukleotids des ersten ligierten Adaptor-Moleküls (s. (4b)) mit zusätzlich drei 3'-Nukleotiden der vollständigen Permutation aller vier Desoxyribo-nukleotide dA, dC, dG und dT entsprechen, verwendet. Die hier eingesetzten Primer-Permutationen entsprechen dabei den in den Ligations-Ansätzen unter (4b) definierten Basen 2-4 der Adaptoren-Überhänge. Als 3'-Primer werden vier Oligonukleotide der Länge 18-27 Basen, die dem 3'-Ende des Sense-Oligonukleotids des zweiten ligierten Adaptor-Moleküls (s. (7)) entsprechen, eingesetzt.

Das PCR-Profil ist typischerweise: 3 min, 95°C gefolgt von 20-40 Zyklen mit 45 sec, 95°C, 45 sec, 65°C, 60 sec, 72°C und einer abschliessenden Extension für 60 sec bei 72°C, und wird mit Hilfe von radioaktiv oder fluoreszierend markierten PCR-Primern durchgeführt.

#### 8d) PCR für die Alternative (4c)

In insgesamt 256 PCRs (für jede 'Auffüllreaktion' eine PCR) wird ein 5'-Primer der Länge 18-27 Basen, die dem 3'-Ende des Sense-Oligonukleotids des ligierten Adaptor-Moleküls entspricht, eingesetzt. Als 3'-Primer wird das cDNA-Synthese-Oligonukleotid verwendet. Das PCR-Profil ist typischerweise: 3 min, 95°C gefolgt von 20-40 Zyklen mit 45 sec, 95°C, 45 sec, 65°C, 60 sec, 72°C und einer abschliessenden Extension für 60 sec bei 72°C, und wird in Gegenwart von radioaktiv markierten Nukleotiden durchgeführt.

#### 9) PCR-Analyse

Die Analyse der PCR-Fragmente erfolgt typischerweise auf 6 %-igen Polyacrylamidgelen (PAA-Gelen) mit 7-8 M Harnstoff oder durch Kapillar-Elektrophorese.

Aus den Ergebnissen zeigt sich, dass die Selektivität für die Amplifikation der korrekten 5'-Nukleotide der cDNA für die 5'-Primer der TOGA-Methode nicht hinreichend gegeben ist, insbesondere wenn die Durchführung einer Computer-gestützten Datenbank-Analyse der Resultate angestrebt wird. Die Fehler-Rate bei der erfindungsgemässen Verwendung der 5'-Primer beträgt  $\leq 5\%$ . Die Fehler-Rate lässt sich durch die oben beschriebene Liagtions-Methode weiter reduzieren auf annähernd 0 %. Angesichts dieser extrem niedrigen Fehler-Rate wird auch die Etablierung einer automatisierten Daten-Analyse ermöglicht.

Patentansprüche

1. Verfahren zur Identifikation und Charakterisierung von mRNA-Molekülen, welches die folgenden Schritte umfasst:
  - (a) Isolierung und Reinigung von PolyA<sup>+</sup>-RNA aus Gewebeproben;
  - (b) Synthese von doppelsträngiger cDNA aus den mRNA-Molekülen;
  - (c) Verkürzen der cDNA durch enzymatischen Verdau mit Restriktionsendonukleasen;
  - (d) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA; und entweder, in einer 1. Alternative,  - (e) Auffüllen der 5'-Überhänge der cDNA mit Desoxyribonukleotiden und Klenow-DNA-Polymerase;
  - (f) selektive Aufreinigung der 3'-Enden der cDNA;
  - (g) Abtrennung der 3'-ständigen Poly-dA-Nukleotide der cDNA durch enzymatischen Verdau mit einer Restriktionsendonuklease;
  - (h) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA;
  - (i) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction);
  - (j) Auftrennung der Amplifikationsprodukte nach ihrer Länge;
  - (k) Analyse der Amplifikationsprodukte;
oder, in einer 2. Alternative,  - (e) selektive Aufreinigung der 3'-Enden der cDNA;
  - (f) Abtrennung der 3'-ständigen Poly-dA-Nukleotide der cDNA durch enzymatischen Verdau mit einer Restriktionsendonuklease;
  - (g) Hybridisierung und Ligation von Adaptor-Molekülen an die geschnittene cDNA;
  - (h) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction), wobei 5'-Primer verwendet werden, die zwei oder drei Permutationen am 3'-Ende aufweisen, und die eine oder zwei artifizielle Fehlpaarungen im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang enthalten;
  - (i) Auftrennung der Amplifikationsprodukte nach ihrer Länge;
  - (j) Analyse der Amplifikationsprodukte;
oder, in einer dritten Alternative,  - (e) Amplifizierung der cDNA-Fragmente durch PCR (polymerase chain reaction), wobei 5'-Primer verwendet werden, die zwei oder drei Permutationen am 3'-Ende

aufweisen, und die eine oder zwei artifizielle Fehlpaarungen im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang enthalten;

(f) Auftrennung der Amplifikationsprodukte nach ihrer Länge;

(g) Analyse der Amplifikationsprodukte;

wobei für die 1. und 2. Alternative in Schritt (b) die Synthese der cDNA-Erststrangmoleküle durch reverse Transkription unter Verwendung eines geankerten Oligo-dT-Nukleotids erfolgt, welches eine 3'-Extension von 2 Basen aufweist, wobei die erste Base dA, dC oder dG, und die zweite Base dA, dC, dG oder dT ist, und welches eine 5'-Extension von 5-15 Basen aufweist, die für die Schnittstelle einer Restriktionsendonuklease mit der Spalt-Charakteristik 16/14 'downstream' der Erkennungssequenz kodiert; oder

wobei für die 3. Alternative in Schritt (b) die Synthese der cDNA-Erststrangmoleküle durch reverse Transkription unter Verwendung eines geankerten Oligo-dT-Nukleotids erfolgt, welches eine 3'-Extension von 2 Basen aufweist, wobei die erste Base dA, dC oder dG, und die zweite Base dA, dC, dG oder dT ist, und welches eine 5'-Extension von 5-15 Basen einer beliebigen Sequenz aufweist.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass das Oligo-dT-Nukleotid vollständig durch 2'-O-methylierte Ribonukleotide substituiert wird.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass das Oligo-dT-Nukleotid an seinem freien 5'-Ende und/oder an internen dT-Nukleotiden mit einem Biotinrest über einen C9-Spacer versehen wird.
4. Verfahren nach Anspruch 1, wobei in Schritt (h) der 2. Alternative oder in Schritt (e) der 3. Alternative sich die Fehlpaarungen an den Positionen -3, oder -3 und -4, oder -4 und -5 im Vergleich zum durch die Adaptermoleküle bestimmten komplementären Strang befinden.
5. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass in Schritt (c) ein Restriktionsenzym der Klasse IIS verwendet wird, welches 5 Nukleotide als Erkennungssequenz besitzt und einen aus 2 - 4 Nukleotiden, die nicht Teil der

Erkennungssequenz sind, bestehenden Überhang der geschnittenen cDNA-Fragmente generiert.

6. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass in Schritt (f) der 1. Alternative oder in Schritt (e) der zweiten Alternative die selektive Aufreinigung der 3'-Enden der cDNA unter Verwendung von paramagnetischen 'beads' erfolgt, die das Biotin-bindende Molekül Streptavidin gekoppelt haben.
7. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass in Schritt (g) der 1. Alternative oder in Schritt (f) der 2. Alternative ein Restriktionsenzym der Klasse IIS verwendet wird.
8. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Produkte aus Schritt (d) oder aus Schritt (h) der 1. Alternative oder aus Schritt (g) der 2. Alternative vor der Amplifizierung gemäss Schritt (i) mit einer Nuklease, ausgewählt aus der Gruppe bestehend aus T4-Endonuclease VII, S1-Nuclease, und Mung Bean Nuclease, inkubiert werden.
9. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass in Schritt (i) der ersten Alternative oder in Schritt (h) der 2. Alternative oder in Schritt (e) der 3. Alternative die Amplifizierung der cDNA-Fragmente unter Verwendung von Oligonukleotiden erfolgt, die an den komplementären Strang des 'sense'-Oligomers der ligierten Adaptor-Moleküle hybridisieren.
10. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass in Schritt (k) der ersten Alternative oder in Schritt (j) der 2. Alternative oder in Schritt (g) der 3. Alternative die Analyse anhand der unterschiedlichen Längen der Produkte sowie in Kenntnis der durch die Manipulation bekannten Basensequenz von 9 oder 10 Nukleotiden erfolgt.
11. Anwendung des Verfahrens nach einem der vorhergehenden Ansprüche zur gegebenenfalls rechnergestützten Identifizierung und Isolierung sowie Analyse neuer Gene.



## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 99/00992

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 05286 A (UNIV YALE ;WEISSMAN SHERMAN (US); PRASHAR YATINDRA (US)) 13 February 1997 see the whole document, in particular claims	1
X	WO 95 13369 A (SCRIPPS RESEARCH INST) 18 May 1995 cited in the application see the whole document, in particular claims	1
Y	WO 94 01582 A (MEDICAL RES COUNCIL ;SIBSON DAVID ROSS (GB)) 20 January 1994 see the whole document, in particular claims (14 ff.) and figures -/-	1-11



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

15 June 1999

Date of mailing of the international search report

22/06/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Müller, F

# INTERNATIONAL SEARCH REPORT

Internat'l Application No  
PCT/EP 99/00992

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LIANG P ET AL: "DIFFERENTIAL DISPLAY OF EUKARYOTIC MESSENGER RNA BY MEANS OF THE POLYMERASE CHAIN REACTION" SCIENCE, vol. 257, no. 5072, 14 August 1992, pages 967-971, XP000508268 cited in the application see the whole document	1-11
X	WO 97 29211 A (US HEALTH ;WEINSTEIN JOHN N (US); BOULAMWINI JOHN (US)) 14 August 1997 cited in the application see claims and figures	1
X	EP 0 743 367 A (MAX PLANCK GESELLSCHAFT) 20 November 1996 claims and figures	1
X	EP 0 735 144 A (JAPAN RES DEV CORP) 2 October 1996 see the whole document	1,4,7,11
A	VELCULESCU V E ET AL: "SERIAL ANALYSIS OF GENE EXPRESSION" SCIENCE, vol. 270, 20 October 1995, pages 484-487, XP002053721 see the whole document	
A	UNRAU P ET AL: "NON-CLONING AMPLIFICATION OF SPECIFIC DNA FRAGMENTS FROM WHOLE GENOMIC DNA DIGEST USING DNA 'INDEXERS'" GENE, vol. 145, 1 January 1994, pages 163-169, XP002054436 see the whole document	
A	WO 97 28282 A (STRATAGENE INC) 7 August 1997 see claims	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Internal Application No

PCT/EP 99/00992

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9705286 A	13-02-1997	US 5712126 A AU 692403 B AU 6763696 A CA 2201468 A EP 0797685 A JP 10509329 T	27-01-1998 04-06-1998 26-02-1997 13-02-1997 01-10-1997 14-09-1998
WO 9513369 A	18-05-1995	US 5459037 A AU 687127 B AU 1055195 A AU 6711098 A CA 2174966 A EP 0726946 A FI 962000 A JP 9509306 T NO 961902 A US 5807680 A	17-10-1995 19-02-1998 29-05-1995 09-07-1998 18-05-1995 21-08-1996 10-05-1996 22-09-1997 12-07-1996 15-09-1998
WO 9401582 A	20-01-1994	AT 159986 T AU 686563 B AU 4575893 A CA 2139944 A DE 69315074 D DE 69315074 T EP 0650528 A JP 7508883 T US 5728524 A	15-11-1997 12-02-1998 31-01-1994 20-01-1994 11-12-1997 05-03-1998 03-05-1995 05-10-1995 17-03-1998
WO 9729211 A	14-08-1997	AU 2264197 A	28-08-1997
EP 0743367 A	20-11-1996	DE 19518505 A JP 8308598 A US 5876932 A	21-11-1996 26-11-1996 02-03-1999
EP 0735144 A	02-10-1996	JP 2763277 B JP 9028399 A JP 2763278 B JP 8322598 A AU 692685 B AU 5031196 A US 5707807 A	11-06-1998 04-02-1997 11-06-1998 10-12-1996 11-06-1998 10-10-1996 13-01-1998
WO 9728282 A	07-08-1997	NONE	

## A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES

IPK 6 C12Q1/68

Nach der internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

## B. RESEARCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 6 C12Q

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

## C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 97 05286 A (UNIV YALE ;WEISSMAN SHERMAN (US); PRASHAR YATINDRA (US)) 13. Februar 1997 siehe ganzes Dokument, insbes. Ansprüche	1
X	WO 95 13369 A (SCRIPPS RESEARCH INST) 18. Mai 1995 in der Anmeldung erwähnt siehe ganzes Dokument, insbes. Ansprüche	1
Y	WO 94 01582 A (MEDICAL RES COUNCIL ;SIBSON DAVID ROSS (GB)) 20. Januar 1994 siehe ganzes Dokument, insbes. Ansprüche (14 ff.) und Abbildungen --- -/-	1-11



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

\* Besondere Kategorien von angegebenen Veröffentlichungen :

\*A\* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

\*E\* älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

\*L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

\*O\* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

\*P\* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

\*T\* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

\*X\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderscher Tätigkeit beruhend betrachtet werden

\*Y\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderscher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

\*Z\* Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

15. Juni 1999

Absenddatum des internationalen Recherchenberichts

22/06/1999

Name und Postanschrift der Internationalen Recherchenbehörde

Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Müller, F

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	<p>LIANG P ET AL: "DIFFERENTIAL DISPLAY OF EUKARYOTIC MESSENGER RNA BY MEANS OF THE POLYMERASE CHAIN REACTION"</p> <p>SCIENCE, Bd. 257, Nr. 5072, 14. August 1992, Seiten 967-971, XP000508268</p> <p>in der Anmeldung erwähnt siehe das ganze Dokument</p>	1-11
X	<p>WO 97 29211 A (US HEALTH ;WEINSTEIN JOHN N (US); BOULAMWINI JOHN (US))</p> <p>14. August 1997</p> <p>in der Anmeldung erwähnt siehe Ansprüche und Abbildungen</p>	1
X	<p>EP 0 743 367 A (MAX PLANCK GESELLSCHAFT)</p> <p>20. November 1996</p> <p>Ansprüche und Abbildungen</p>	1
X	<p>EP 0 735 144 A (JAPAN RES DEV CORP.)</p> <p>2. Oktober 1996</p> <p>siehe das ganze Dokument</p>	1,4,7,11
A	<p>VELCULESCU V E ET AL: "SERIAL ANALYSIS OF GENE EXPRESSION"</p> <p>SCIENCE, Bd. 270, 20. Oktober 1995, Seiten 484-487, XP002053721</p> <p>siehe das ganze Dokument</p>	
A	<p>UNRAU P ET AL: "NON-CLONING AMPLIFICATION OF SPECIFIC DNA FRAGMENTS FROM WHOLE GENOMIC DNA DIGEST USING DNA 'INDEXERS'"</p> <p>GENE, Bd. 145, 1. Januar 1994, Seiten 163-169, XP002054436</p> <p>siehe das ganze Dokument</p>	
A	<p>WO 97 28282 A (STRATAGENE INC)</p> <p>7. August 1997</p> <p>siehe Ansprüche</p>	

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

internationalis Aktenzeichen

PCT/EP 99/00992

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9705286 A	13-02-1997	US 5712126 A AU 692403 B AU 6763696 A CA 2201468 A EP 0797685 A JP 10509329 T	27-01-1998 04-06-1998 26-02-1997 13-02-1997 01-10-1997 14-09-1998
WO 9513369 A	18-05-1995	US 5459037 A AU 687127 B AU 1055195 A AU 6711098 A CA 2174966 A EP 0726946 A FI 962000 A JP 9509306 T NO 961902 A US 5807680 A	17-10-1995 19-02-1998 29-05-1995 09-07-1998 18-05-1995 21-08-1996 10-05-1996 22-09-1997 12-07-1996 15-09-1998
WO 9401582 A	20-01-1994	AT 159986 T AU 686563 B AU 4575893 A CA 2139944 A DE 69315074 D DE 69315074 T EP 0650528 A JP 7508883 T US 5728524 A	15-11-1997 12-02-1998 31-01-1994 20-01-1994 11-12-1997 05-03-1998 03-05-1995 05-10-1995 17-03-1998
WO 9729211 A	14-08-1997	AU 2264197 A	28-08-1997
EP 0743367 A	20-11-1996	DE 19518505 A JP 8308598 A US 5876932 A	21-11-1996 26-11-1996 02-03-1999
EP 0735144 A	02-10-1996	JP 2763277 B JP 9028399 A JP 2763278 B JP 8322598 A AU 692685 B AU 5031196 A US 5707807 A	11-06-1998 04-02-1997 11-06-1998 10-12-1996 11-06-1998 10-10-1996 13-01-1998
WO 9728282 A	07-08-1997	KEINE	